Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Distributed Image and Video Coding

based on

Compressed Sensing

A thesis presented in partial fulfilment of the requirements for the degree of

Doctor of Philosophy in Computer Engineering at Massey University New Zealand.

Muhammad Yousuf Baig

2013

Abstract

Conventional methods for encoding of images and videos is a complex process with high computational demands. They are designed for application scenarios where the signals concerned are encoded once and played back many times. However, new applications such as wireless video sensor networks demand low cost and low power cameras with limited computing resources. The focus of this thesis is on such image and video coding systems where the computational burden is shifted from the encoder to the decoder.

Three separate coding schemes have been developed – two for videos and one for images. Together they form a framework for distributed coding which is based on the theory of compressed sensing and distributed coding. Compressed sensing is a relatively new theory for the acquisition of sparse signals that allows the sampling rate to be much lower than the Nyquist limit. Distributed coding is based on the theorem by Slepian and Wolf, and Wyner and Ziv. It allows different correlated parts of a signal to be encoded independently without loss of coding efficiency. The decoding of these separately encoded parts are then decoded jointly in order to exploit the correlation between them. The main characteristics of the coding scheme proposed in this thesis are: (1) they do not require the use of traditional codecs; (2) only compressed sensing measurements are used for encoding and decoding; (3) no motion estimation and compensation are involved for videos.

The first proposed coding scheme is for the encoding of whole video frames. The compressed sensing measurement of individual frames are separately encoded. These frames are divided into key and non-key frames with the key frames encoded at a higher rate than non-key ones. While the key frames are decoded independently, the non-key ones are decoded with the help of side information generated from the measurements of the key frames. The most important part of the decoder is a simple, yet effective, side information generation method which requires only minimal computation. The side information generated is simply added to the measurements of the non-key frames for use with any compressed sensing reconstruction algorithm. The other two coding schemes are block-based coding methods. Each image or frame is divided into non-overlapping image blocks in a similar way it is done in some existing coding standards. The coding of the blocks are performed in a distributed manner by classifying them into key blocks and nonkey blocks. An adaptive encoding strategy based on block similarity is also developed. Experimental analyses using publicly available test images and videos show that the performances of the simpler codecs proposed are better than other existing compressed sensing based codecs. The video codecs also out-perform conventional distributed video codec in terms of simplicity, compression ratio and decoding complexity.

The basis of these coding methods is on the correlation of frames or blocks. This correlation is established through experimental analyses. These analyses also showed that the minimum square error between any pair of them can be effectively used as a measure of correlation. In conjunction with the development of the codecs, a quantization scheme that is tailored to the statistics of CS measurements has also been proposed. This scheme yields better results than a uniform quantizer and those used for JPEG. The quantizer is also robust against different statistics of individual images. Separate experimental evaluations also show that structurally random matrices are the best sensing matrices for acquiring images and the sparse reconstruction by separable approximation (SpaRSA) algorithm produces the best reconstructed image quality.

Acknowledgments

I am profoundly grateful to my supervisor Associate Professor Edmund Lai for his constant support and guidance throughout my PhD study. There were many occasion in this journey when things were not going in a right direction but Edmund's encouragements and endless patience has helped me to finish this thesis. Edmund's knowledge and inspiring research insights were always fruitful and without him, this endeavour could have never been completed.

I am also thankful to my co supervisor Dr Amal Punchihewa for his valuable comments, suggestions and feedback throughout my research. This PhD thesis could not be completed without a financial support and scholarship from Higher Education Commission, Pakistan. I am thankful for their support and appreciate their effort for educational reforms. I am also thankful to Mehran University for granting me study leave to pursue PhD study.

I want to express my gratitude to my parents for their prayers and support. I am thankful to my wife Sana for her love, understanding and endless support despite all the difficulties. Finally love to my children, Maryam and Taha who always brings smiles back in a long and hard day

Contents

	Abs	stract	i
	Ack	knowledgments	iii
	Lis	st of Figures	х
	List of Tables		
	Lis	st of Abbreviations	
			XV
1	Intr	roduction	1
	1.1	Background and Motivation	1
	1.2	Scope and Objectives	3
	1.3	Original Contributions	4
	1.4	Organization of the Thesis	5
2	An	Overview of Compressed Sensing	7
	2.1	Key Elements of Compressed Sensing	7
		2.1.1 Sparsity	8
		2.1.2 Incoherence	10
		2.1.3 CS Measurement Acquisition	10

		2.1.4	CS Reconstruction	13
	2.2	Poten	tial Applications	15
	2.3	Summ	nary	18
3	A R	Review	of Compressed Sensing Image and Video Coding	19
	3.1	Comp	ressed Sensing Image Coding	19
		3.1.1	Full Image Coding	19
		3.1.2	Block Based Coding	20
		3.1.3	Multi-scale Coding	22
		3.1.4	Distributed Coding	24
	3.2	Comp	ressed Sensing Video Coding	29
		3.2.1	3D Transform Coding	29
		3.2.2	Distributed Coding	31
		3.2.3	Dictionary-based Coding	35
		3.2.4	Residue-based Coding	37
		3.2.5	Adaptive Coding	41
		3.2.6	Scalable Coding	45
	3.3	Conve	entional Distributed Video Coding	46
		3.3.1	Transform Domain Wyner-Ziv Video Coding	47
		3.3.2	The PRISM Video Codec	47
		3.3.3	The DISCOVER Video Codec	49
		3.3.4	Comparison of DVC Architectures	50
	3.4	Summ	arv	52

4	Sen	$\operatorname{sing} M$	latrix, Quantization Matrix and Reconstruction Algorith	\mathbf{ms}	
	for Image Compression		5	3	
	4.1	Choice	e of Sensing Matrices	. 5	5
		4.1.1	Experimental Results	. 5	6
	4.2	Choice	e of Reconstruction Algorithms	. 6	1
		4.2.1	Greedy Algorithms	. 6	3
		4.2.2	Gradient based Algorithms	. 6	4
		4.2.3	Iterative Shrinkage Thresholding Algorithms	. 6	7
		4.2.4	Experimental Results	. 6	8
	4.3	Design	of Quantizer	. 7	2
		4.3.1	Distribution of CS Measurements	. 7	3
		4.3.2	Proposed Quantization Scheme	. 7	5
	4.4	Summ	ary	. 7	7
5	Dist	tribute	d Inter-frame Video Compressed Sensing	7	9
	5.1	Propos	sed DCVS Codec	. 8	0
		5.1.1	Encoder	. 8	1
		5.1.2	Decoder	. 8	2
		5.1.3	Correlation Analysis of CS Measurements	. 8	3
		5.1.4	Correlation and Mean Square Error	. 8	6
	5.2	Side In	nformation Generation	. 8	6
		5.2.1	Motion Compensated Interpolation	. 8	8
		5.2.2	Proposed SI Generation Method	. 9	0

vii

5.3	CS Reconstruction with Side Information
5.4	Experimental Results
	5.4.1 Reconstruction Complexity Evaluation
	5.4.2 Rate Distortion Evaluation $\dots \dots \dots$
	5.4.3 Performance Comparison with Distributed and Conventional Codecs103
5.5	Summary
Dist	ributed CS Image Compression 113
6.1	Block Based Encoding and Decoding
	6.1.1 Recovery Methods \ldots 114
	6.1.2 Impact of Block Size
6.2	Block Similarity Analysis
6.3	Proposed Distributed Image Codec
	$6.3.1 \text{Encoder} \dots \dots \dots \dots \dots \dots \dots \dots \dots $
	$6.3.2 \text{Decoder} \dots \dots \dots \dots \dots \dots \dots \dots \dots $
6.4	Experimental Results
6.5	Summary
Dist	ributed Block-based Video Compressed Sensing 133
7.1	Block Correlation Analysis
	7.1.1 Intra-frame Block Correlation
	7.1.2 Inter-frame Block Correlation
7.2	Proposed Adaptive Block-based Video Codec
	 5.3 5.4 5.5 Distr 6.1 6.2 6.3 6.4 6.5 Distr 7.1 7.2

		7.2.1	Encoder	140
		7.2.2	Decoder	142
	7.3	Experi	mental Results	143
		7.3.1	Measurement Rate Reduction	146
		7.3.2	Reconstruction Complexity Evaluation	146
		7.3.3	Rate Distortion Evaluation	148
		7.3.4	Performance Comparison with DISCOVER and Other Conven-	
			tional Codecs	154
	7.4	Summ	ary	162
	1.1			
8	Con	clusio	ns and Future Work	167
8	Con 8.1	clusio Conclu	ns and Future Work	167 167
8	Con 8.1 8.2	clusio Conclu Furthe	ns and Future Work	167 167 169
8	Con 8.1 8.2	clusion Conclu Furthe 8.2.1	ns and Future Work usions	167167169169
8	Con 8.1 8.2	clusion Conclu Furthe 8.2.1 8.2.2	ns and Future Work usions or Research Multi-view Image/Video Coding Hyperspectral Imaging	 167 167 169 169 169
8 A	Con 8.1 8.2 List	clusion Conclu Furthe 8.2.1 8.2.2 of Pu	ns and Future Work Isions	 167 167 169 169 169 171

List of Figures

2.1	Key Elements of Compressed Sensing	9
2.2	CS Camera, [Image Courtesy, InView Corporation [41]]	16
3.1	Multi Scale Image Coding Scheme [59]	23
3.2	Wyner-Ziv Image Coding Scheme [68]	25
3.3	Wyner-Ziv Image Coding Architecture [69].	26
3.4	Architecture of Adaptive Distributed Image Sensing [71]	28
3.5	Architecture of Distributed Compressed Video Sensing [13]	31
3.6	Distributed CS Encoder / Decoder [14]	33
3.7	CS based Video Coder [17]. \ldots \ldots \ldots \ldots \ldots \ldots \ldots	36
3.8	Residual CS Encoder / Decoder [89]	39
3.9	k-t FOCUSS with ME/MC [78]	40
3.10	Block based Video Codec [15]	42
3.11	Adaptive Block based Video Codec [79].	43
3.12	Transform Domain Wyner-Ziv video Coding architecture [99]	46
3.13	PRISM Video Coding Architecture [101]	48
3.14	Block diagram of the DISCOVER video coding architecture [103]	49

4.1	CS Image Compression	54
4.2	Sensing Matrix Acquisition Time	57
4.3	Sensing Matrix Reconstruction Time	59
4.4	Sensing Matrix Rate Distortion Performance	60
4.5	Reconstruction Times for CS Reconstruction Algorithms	69
4.6	Number of Iterations Required for CS Reconstruction	70
4.7	Rate-Distortion Performance of CS Reconstruction Algorithms	71
4.8	CS Measurements and DCT Coefficients Histogram for Test Images $\ . \ .$.	74
4.9	Quantization Matrix Rate-Distortion Performance	76
4.10	Reconstruction visual quality for Lena	78
5.1	Proposed Video Codec	81
5.2	Correlation Analysis for CS Measurements	84
5.3	Correlation of CS Measurements of WZ frames with Key frames	85
5.4	Correlation and MSE Comparison of CS Measurements of WZ frames with	
	Key frames	87
5.5	DVC Side Information Generation	89
5.6	Motion Compensated Interpolation	90
5.7	Median of Laplacian Distribution Parameter for Two Types of SI $\ . \ . \ .$	92
5.8	Reconstruction complexity comparison of Video Sequences $\ldots \ldots \ldots$	97
5.9	Rate Distortion Curve for GOP Size 3	103
5.10	Rate Distortion Curve for GOP Size 5	104
5.11	Rate Distortion Curve for GOP Size 8	105

5.12	Visual Reconstruction Quality of News 89th Frame for GOP Size 3	106
5.13	Visual Reconstruction Quality of Container 56th Frame for GOP Size 3 .	107
5.14	Bit Rate vs Compression Ratio for Video Sequences - GOP Size 3 $\ . \ . \ .$	109
5.15	Bit Rate vs PSNR for Video Sequences - GOP Size 3	110
5.16	Reconstruction Time Complexity Comparison with DISCOVER $\ . \ . \ .$.	111
6.1	Block Based CS Reconstruction Comparison	115
6.2	Rate Distortion Performance of Block Size (64, 32, 16, and 8) \ldots	117
6.3	Block Similarity Analysis of Original Pixel Data and CS measurements with Correlation Coefficient and MSE	119
6.4	Percentage of Similar Blocks for Block Size (64, 32, 16 and 8)	120
6.5	Proposed Distributed Intra Image Codec	122
6.6	Test Images used	126
6.7	Rate-Distortion Performance for Test Images	128
6.8	SSIM Index for Test Images	129
6.9	Compression Efficiency for Test Images	130
7.1	Intra Block Correlation of Original Pixel Data and CS measurements with MSE and Correlation Coefficient	135
7.2	Percentage of Similar Blocks for Block Size (64, 32, 16 and 8) \ldots	136
7.3	Inter Block Correlation of Original Pixel Data and CS measurements be- tween 1st and 2nd frame of Foreman video with MSE and Correlation	190
	Coemcient	138

7.4	Inter Block Correlation of Original Pixel Data and CS measurements be- tween 1st and 3rd frame of Foreman video with MSE and Correlation	
	Coefficient	138
7.5	Inter Block Correlation of Original Pixel Data and CS measurements be- tween 1st and 2nd frame of Coastguard video with MSE and Correlation Coefficient	139
7.6	Inter Block Correlation of Original Pixel Data and CS measurements be- tween 1st and 3rd frame of Coastguard video with MSE and Correlation Coefficient	139
7.7	Proposed Distributed Block-based Video Codec	141
7.8	Test Videos	144
7.9	Reconstruction complexity comparison for GOP Size 3, 5 and 8 $\ .$	150
7.10	Rate Distortion Curve for GOP Size 3	151
7.11	Rate Distortion Curve for GOP Size 5	153
7.12	Rate Distortion Curve for GOP Size 8	154
7.13	SSIM Index for GOP Size 3	156
7.14	SSIM Index for GOP Size 5	157
7.15	SSIM Index for GOP Size 8	158
7.16	Visual Reconstruction Quality of Akiyo 123rd Frame for GOP Size 3 $$	159
7.17	Visual Reconstruction Quality of Mother Daughter 24th Frame for GOP Size 3	160
7.18	Bit Rate vs Compression Ratio for Video Sequences - GOP Size 3 $\ . \ . \ .$	162
7.19	Bit Rate vs PSNR for Video Sequences - GOP Size 3	163
7.20	Reconstruction Time Complexity Comparison with DISCOVER	164

List of Tables

4.1	Performance Comparison of Quantization Schemes, in PSNR(dB)	77
5.1	Video Test Sequences	94
5.2	Average Reconstruction Time (in Seconds) of Video Sequences	96
5.3	Rate Distortion Performance (in dB) of Video Sequences for GOP Size 3	99
5.4	Rate Distortion Performance (in dB) of Video Sequences for GOP Size 5 $$	100
5.5	Rate Distortion Performance (in dB) of Video Sequences for GOP Size 8	101
5.6	Average SSIM Index Performance of Video Sequences for GOP Size 3, 5 $$	
	and 8	102
6.1	Performance Evaluation of Test Images	127
7.1	Percentage of Measurement Rate Reduction in Adaptive Encoding	147
7.2	Average Reconstruction Time (in Seconds) of Video Sequences	149
7.3	Average Rate Distortion Performance of Video Sequences for GOP Size 3,	
	5 and 8	152
7.4	Average SSIM Index Performance of Video Sequences for GOP Size 3, 5 $$	
	and 8	155

List of Abbreviations

AVC	Advanced Video Coding
BCH Codes	Bose Chaudhuri Hocquenghem Codes
BCQP	Bounded Constraint Quadratic Program
BCS	Bayesian Compressive Sensing
BM	Block Matching
BP	Basis Pursuit
BPDN	Basis Pursuit Denoising
CoSaMP	Compressive Sampling Matching Pursuit
\mathbf{CRC}	Cyclic Redundancy Check
\mathbf{CS}	Compressed Sensing
DCT	Discrete Cosine Transform
DCVS	distributed compressed video coding
DISCOS	Distributed Compressed Video Sensing
DISCOVER	DistributedCoding forVideo Services
\mathbf{DSC}	Distributed Source Coding
DVC	Distributed Video Coding
\mathbf{Ffmpeg}	Fast Foraward MPEG
GOB	Group of Blocks
GOP	Group of Picture
GPSR	Gradient Projection for Sparse Reconstruction
ISAR	Inverse Synthetic Aperture Radar
ITU-T	International Telecommunication Union
JPEG	Joint Picture Expert Group
k-t FOCUSS	Focal Under Determined System Solver in k-t Space
LASSO	Least Aboslute Shrinkage and Selection Operator
LDPC	Low Density Parity Check
LDPCA	Low Density Parity Check Accumulate
LIMAT	Lifting-based Invertible Motion Adaptive Transform
LP	Linear Programming
\mathbf{MC}	Motion Compensation
MCFI	Motion Compensated Frame Interpolation
MCI	Motion Compensated Interpolation
\mathbf{ME}	Motion Estimation
\mathbf{MP}	Matching Pursuit
MPEG	Motion Picture Expert Group
\mathbf{MR}	Measurement Rate
MRI	Magnetic Resonance Imaging
MSE	Mean Square Error
NESTA	Nesterov's Algorithm

NP	Non-deterministic Polynomial-time
OMP	Orthogonal Matching Pursuit
PCA	Principle Component Analysis
\mathbf{PFT}	Partial Fourier Transform
PL	Projected Landweber
PRISM	Power-efficient, Robust, hIgn compression Syndrome
	based Multimedia coding
PSNR	Peak Signal to Noise Ratio
QCIF	Quad Common Interchange Format
RD	Rate Distortion
RDO	Rate Distortion Optimization
RIP	Restricted Isometry Property
SBHE	Scrambled Block Hadamard Ensemble
SFE	Scrambled Fourier Ensemble
SI	Side Information
SpaRSA	Sparse Reconstruction by Seperable Approximation
\mathbf{SPL}	Smoothed Projected Landweber
\mathbf{SRM}	Structurally Random Matrices
SSIM	Structure Similarity Index
STD	Standard Deviation
StOMP	Stagewise Orthogonal Matching Pursuit
SVC	Scalable Video Coding
SVD	Singular Value Decomposition
\mathbf{TV}	Total Variation
TwIST	Two-Step Iterative Shrinkage Thresholding
TWR	Through-the-wall Radar
\mathbf{WT}	Wavelet Transform
WZ	Wyner Ziv